ehp ENVIRONMENTAL HEALTH PERSPECTIVES
http://www.ehponline.org

Sex-Specific Associations between Particulate Matter Exposure and Gene Expression in Independent Discovery and Validation Cohorts of Middle-Aged Men and Women

Karen Vrijens, Ellen Winckelmans, Maria Tsamou, Willy Baeyens, Patrick De Boever, Danyel Jennen, Theo M. de Kok, Elly Den Hond, Wouter Lefebvre, Michelle Plusquin, Hans Reynders, Greet Schoeters, Nicolas Van Larebeke, Charlotte Vanpoucke, Jos Kleinjans, and Tim S Nawrot

http://dx.doi.org/10.1289/EHP370

Received: 21 December 2015
Revised: 12 August 2016
Accepted: 22 August 2016
Published: 14 October 2016

Environ Health Perspect DOI: 10.1289/EHP370
Advance Publication: Not Copyedited

Sex-Specific Associations between Particulate Matter Exposure and Gene Expression in Independent Discovery and Validation Cohorts of Middle-Aged Men and Women

Karen Vrijens[1], Ellen Winckelmans[1], Maria Tsamou[1], Willy Baeyens[2], Patrick De Boever[1,3], Danyel Jennen[4], Theo M. de Kok[4], Elly Den Hond[3,5], Wouter Lefebvre[3], Michelle Plusquin[1], Hans Reynders[6], Greet Schoeters[3,7,8], Nicolas Van Larebeke[9], Charlotte Vanpoucke[10], Jos Kleinjans[4], and Tim S Nawrot[1,11]

[1]Centre for Environmental Sciences, Hasselt University, Agoralaan gebouw D, Diepenbeek, 3590, Belgium, [2]Department of Analytical and Environmental Chemistry, Free University of Brussels, Brussels, Belgium, [3]Environmental Risk and Health, Flemish Institute for Technological Research Mol, Belgium, [4]Department of Toxicogenomics, Maastricht University, Maastricht, The Netherlands, [5]Provincial Institute for Hygiene, Antwerp, Belgium. [6]Flemish Government, Environment, Nature and Energy Department, Belgium, [7]Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium; [8]University of Southern Denmark, Institute of Public Health, Department of Environmental Medicine, Odense, Denmark; [9]Department of Radiotherapy and Nuclear Medicine, Ghent University, Ghent, Belgium; [10]Belgian Interregional Environment Agency (IRCEL), Brussels, Belgium; [11]Department of Public Health & Primary Care, Leuven University, Leuven, Belgium

Corresponding author: Tim S Nawrot
Centre for Environmental Sciences
Hasselt University
Agoralaan gebouw D
B-3590 Diepenbeek
Belgium
Tel: 0032/11-26.83.82
Email: tim.nawrot@uhasselt.be

Competing interests

The authors declare they have no competing financial interests.

Funding

The project was funded by the Environment, Nature and Energy Department of the Flemish government (LNE/OL201100023/13034/M&G), Steunpunt Milieu- en Gezondheid and European Research Council (ERC-2012-StG 310898). Karen Vrijens is a postdoctoral Fellow of the Research Foundation –Flanders (12D7714N).

Running title: Transcriptomic markers of air pollution exposure

Abstract

Background: Particulate matter (PM) exposure leads to premature death, mainly due to respiratory and cardiovascular diseases.

Objectives: Identification of transcriptomic biomarkers of air pollution exposure and effect in a healthy adult population.

Methods: Microarray analyses were performed in 98 healthy volunteers (48 men, 50 women). The expression of 8 sex-specific candidate biomarker genes (significantly associated with PM_{10} in the discovery cohort and with a reported link to air pollution-related disease) was measured with qPCR in an independent validation cohort (75 men, 94 women). Pathway analysis was performed using Gene Set Enrichment Analysis. Average daily $PM_{2.5}$ and PM_{10} exposures over 2-years were estimated for each participant's residential address using spatiotemporal interpolation in combination with a dispersion model

Results: Average long-term PM_{10} was 25.9 (± 5.4) and 23.7 (±2.3) $\mu g/m^3$ in the discovery and validation cohorts, respectively. In discovery analysis, associations between PM_{10} and the expression of individual genes differed by sex. In the validation cohort, long-term PM_{10} was associated with the expression of *DNAJB5* and *EAPP* in men and *ARHGAP4* (p=0.053) in women. *AKAP6* and *LIMK1* were significantly associated with PM_{10} in women, although associations differed in direction between the discovery and validation cohorts. Expression of the 8 candidate genes in the discovery cohort differentiated between validation cohort participants with high *vs* low PM_{10} exposure (area under the receiver operating curve = 0.92; 95% CI: 0.85, 1.00; p=0.0002) in men, 0.86; 95% CI: 0.76, 0.96; p=0.004 in women).

Conclusions: Expression of the sex-specific candidate genes identified in the discovery population predicted PM_{10} exposure in an independent cohort of adults from the same area. Confirmation in other populations may further support this as a new approach for exposure

assessment, and may contribute to the discovery of molecular mechanisms for PM-induced health effects.

Background

Particulate matter (PM) is a complex mixture of small particles and liquid droplets that contains a number of components, including acids, organic chemicals, metals, and soil or dust particles. PM exposure is known to increase overall mortality and morbidity, mainly due to its effect on the cardiorespiratory system (Alfaro-Moreno et al. 2007; Pope et al. 2004). Exposure to PM may disturb normal physiological pathways that maintain homeostasis and this may activate cellular processes that mediate the adverse effects of PM (Kleensang et al. 2014). Gene expression changes play an important role in the activation of pathways of toxicity and gene signatures have the potential to serve as biomarkers of exposure (van Leeuwen et al. 2008; van Breda et al. 2015) and recent reports demonstrate their potential use as biomarkers of effect (La Rocca et al. 2014; Fink et al. 2014). As it has been shown previously that transcriptomic responses to diverse environmental stimuli (i.e. chemical exposure, smoking etc.) can be significantly different between men and women (De Coster et al. 2013; Paul and Amundson 2014), we have opted to perform a sex-specific analysis.

Several studies have suggested that elevated oxidative stress may mediate toxic effects of air pollutants (Donaldson et al. 2005; Nel et al. 2001). The systemic inflammatory response following acute inhalation exposure to PM can induce leukocytosis and monocyte release from the bone marrow (Fujii et al. 2002). Controlled exposure studies of recent diesel exhaust exposure (Pettit et al. 2012) and recent exposure to ultra-fine particles (Huang et al. 2010) have reported evidence of altered gene expression in leukocytes but, to our knowledge, associations between patterns of gene expression and long-term particulate air pollution have not been studied in general populations.

Materials and methods

Study design

As our goal was to identify transcriptomic biomarkers of exposure and effect in a healthy adult population, we started by applying microarray analysis in a discovery cohort of 98 adults for which we modelled particulate matter exposure. On the resulting dataset containing significantly modulated genes and pathways, we applied a literature and bio-informatics approach to identify potential exposure effect biomarkers. Subsequently, these were validated using qPCR analysis in an independent cohort with similar characteristics as the discovery population (Figure 1). Study protocols for the discovery and validation cohort were approved by the Institutional Review Board/Ethical Committee of Antwerp University, and informed consent was obtained from all participants.

Study population

Discovery cohort

The original study population was described previously (van Leeuwen et al. 2008) and consisted of 398 subjects from eight different regions of residence in Flanders (Belgium), as part of the first Flemish Environment and Health Survey (FLEHS I) during the period 2001-2006. Participants were recruited in several communities based on random sampling. Inclusion criteria were age 50–65 years, living in Flanders > 5 years, and being able to complete questionnaires in Dutch. Prior to blood collection, informed consent was obtained from all individuals. A subset of 98 samples was selected for microarray analysis based on previously measured exposure levels to several pollutants including cadmium, lead, PCBs (138, 153 and 180), dioxins, polycyclic aromatic hydrocarbons (PAHs) and benzene. The overall exposure to these pollutants was estimated using a z-score for each pollutant, and study subjects with both low and high exposure levels were chosen for inclusion. Z-scores were not correlated with long-term PM_{10} exposure (r^2=0.0012). Smokers were excluded from

the study population. Paxgene tubes (PreAnalytiX GmbH, Hombrechtikon, Switzerland) were used for RNA collection.

Validation cohort

The qPCR validation study was performed in an independent cohort of 175 adults being part of the third Flemish Environment and Health Survey (FLEHS III) during the period 2012–2015. Healthy volunteers between 50 and 65 years of age, living at the same residential address for at least 10 years and being able to complete questionnaires in Dutch were recruited through registers of general medical practices. Prior to blood collection, informed consent was obtained from all individuals. Participants completed a questionnaire covering age, sex, and smoking habits, among other items, they donated blood and urine samples and subclinical measurements including height, weight and blood pressure were determined. The sampling campaign lasted from May 2014 until the end of the year 2014. We used Paxgene tubes (PreAnalytiX GmbH, Hombrechtikon, Switzerland) to stabilize whole blood RNA for storage.

Exposure estimates

The PM_{10} and $PM_{2.5}$ concentrations for participants' residential addresses were calculated using a spatial temporal interpolation method (Kriging) that takes into account land cover data from satellite images (Corine land cover data set) for interpolating the measurement data of the monitoring stations from the Belgian telemetric air quality network as described previously (Maiheu B et al. 2013; Jacobs et al. 2010; Janssen et al. 2008). Validation statistics of the interpolation tool gave a temporal explained variance of more than 0.7 for hourly PM_{10} averages as well as for annual mean PM_{10} (Maiheu et al. 2013). In combination with a dispersion model (IFDM) using emissions from line sources and point sources, the model chain provides daily PM_{10} and $PM_{2.5}$ values on a 25X 25m receptor grid (Lefebvre et al. 2013). Our model is based on input data from 38 monitoring stations in the study area. The

Initiative on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Use in Europe was the incentive for intensive model intercomparison. IFDM was thoroughly compared with other models currently in use for regulatory purposes in Europe (Olesen,1995; Maes et al. 1995; Cosemans et al. 1995; Mensink et al. 1996; Cosemans et al. 2001).

Mean daily temperatures and relative humidity for the study region were provided by the Royal Meteorological Institute (Brussels, Belgium). Apparent temperature was calculated (Steadman 1979; Kalkstein and Valimont 1986).

All our estimates were annual mean exposures over a 2 year-period because we were interested in developing biomarkers for long-term exposure. For the discovery cohort, annual means were based on 2011-2012 as these were the earliest years for which detailed 25 X 25 m grid information became available. Distribution patterns were used for the year 2008. We assumed that relative differences in annual mean concentrations of particulate matter were generally consistent from year to year. For the validation cohort, annual means were based on the 2 years prior to blood sampling (i.e. 2012-2013).

RNA isolation

Total RNA was isolated from 2.5 mL whole blood from Paxgene Blood RNA vacutainers using the Paxgene Blood RNA system (PreAnalytiX, Qiagen, Hilden, Germany), according to the manufacturer's instructions. A globin reduction assay (GLOBINclear™ Kit by Ambion, Austin, USA) was performed in order to remove hemoglobin mRNA from samples that were submitted to microarray analysis. RNA integrity was assessed using the BioAnalyzer (Agilent, Palo Alto, USA) and purity was measured spectrophotometrically. Labeled samples were checked for specific activity and dye incorporation.

Microarray preparation and hybridization

0.2 µg total RNA from each sample was used to synthesize dye-labeled cRNA (Cy3) following the Agilent one-color Quick-Amp labeling protocol (Agilent Technologies).

Individual samples were hybridized on Agilent 4x44K Whole Human Genome microarrays (design ID 014850).

Microarray data analysis

Microarrays were scanned on an Agilent G2505C DNA Microarray Scanner (Agilent Technologies, Amstelveen, The Netherlands). Raw data on pixel intensities were extracted from the scan images using Agilent Feature Extraction Software (Version 10.7.3.1, Agilent Technologies, Amstelveen, The Netherlands), protocol GE1_107.sep09. Raw data were pre-processed using an in-house developed quality control pipeline in R as follows: local background correction, flagging of bad spots, controls and spots with intensities below background, log2 transformation and quantile normalization. The R-scripts of the pipeline and additional information on the flagging can be found at https://github.com/BiGCAT-UM/arrayQC_Module. From the processed data-files genes were omitted showing more than 30% flagged data, after which the data-files were transferred to the Gene Expression Pattern Analysis Suite, GEPAS 2010 (Montaner et al. 2006) for further pre-processing, including merging replicate probes (based on median), and imputing missing values by means of K-nearest neighbor imputation (K=15). Filtering for flat peaks was used with root mean square value 0.25. The filtered data, containing 28,786 genes were used for further statistical analyses. Microarray gene expression data were analyzed stratified for sex. In the original microarray data set initially 28,786 unique Agilent probe IDs (out of 43,376 Agilent probe IDs) were annotated to 22,390 EntrezGene IDs. In case of multiple replicates (i.e. multiple probes for the same gene), the replicate with highest interquartile range (IQR) in relative gene expression was selected. This resulted in 15,589 unique EntrezGene IDs.

Gene expression analysis

Using linear regression models adjusted for age, body mass index (BMI), socio-economic status (SES, classified in 3 groups: no high school degree, high school degree, or further

education degree), daytime and season of blood sampling we obtained estimates for each gene as the $\log_2$ fold change in gene expression for an increment of 5 $\mu g/m^3$ in exposure. P-values <0.05 were considered statistically significant. P-values were corrected for multiplicity using the Benjamini-Hochberg false discovery rate (FDR) correction. P-values corrected for multiple testing are referred to as q-values.

Pathway analysis

Gene Set Enrichment Analysis was performed utilizing the online pathway analysis tool Consensus Pathway Data Base (CPDB) (http://consensuspathdb.org/). CPDB contains ~5,200 pathways including protein complexes, metabolic, signaling and gene regulatory pathways as well as drug-target interactions. Data originate from 32 public resources curated from the literature (Kamburov et al. 2012). Gene Set Enrichment Analysis was performed in a sex-specific manner using the $\log_2$ fold changes of the gene expression data for all genes analyzed at the gene expression level, without pre-selection. For every predefined gene set in each pathway, a Wilcoxon signed-rank test was calculated, testing the null hypothesis that the distribution of their fold changes was around zero. As input, all genes without *a priori* selection (EntrezGene IDs) were uploaded with their fold changes in their gene expression. We selected the biological processes using pathways as output. The p-values were corrected for multiplicity and were presented as q-values. We defined significant biological processes and pathways by a threshold on the adjusted p-value (q<0.05 or FDR 5%) and we included gene sets with a size between 5 and 100 members.

Selection of potential exposure/effect biomarker genes

We used a modified version of the meet-in-the-middle approach for biomarker identification in relation to clinical relevance, (Vineis et al. 2013) a schematic representation is shown in Figure 1. We first identified the top 50 genes associated with PM_{10} (i.e., with the smallest uncorrected p-values) in men and women, respectively, then performed a literature search

using PubMed and ScienceDirect to identify genes within each sex-specific set that have been associated with air pollution-related health outcomes. Specifically, we searched for the name of each gene in combination with any of the following diseases or processes: allergy (Magnussen et al. 1993), chronic obstructive pulmonary disease (COPD) (Ko et al. 2007), asthma (Bowatte et al. 2015), lung cancer (Raaschou-Nielsen et al. 2011), cardiovascular disease (CVD) (Mills et al. 2009), cerebrovascular disease (CeVD) (Johnson et al. 2010), Alzheimer's disease (Finkelstein and Jerrett) and cognition (Dadvand et al. 2015). Genes with lowest p-values and proven link to AP-related disease were chosen for validation. For men, *DNAJB5*, *RAC3*, *EAPP*, *HDLBP*, *PRG2*, *PER1*, *PIK3R1* and *SLA2* were selected for validation whereas for women the gene list for validation included genes *AKAP6, LIMK1, SIRT7, ARHPGAP4, ATG16L2, TPM3, 5-HTR1B* and *PYGO2*.

Validation of candidate biomarker genes by qPCR

qPCR

Total RNA was reverse transcribed into cDNA by means of the GoScript Reverse Transcription System (Promega, Madison, WI, USA) using a Veriti 96 well Thermal cycler (TC-5000, Techne, Burlington, NJ, USA). A maximum of 3 µg of total RNA was used as input and we used the protocol with an equal amount of oligo(dT) and random hexamer primers according to the manufacturer's instructions. cDNA was stored at -20°C until further measurements. A quantitative real-time polymerase chain reaction (qPCR) was set up by adding 2 µL of a 10 ng/µL dilution of cDNA together with TaqMan Fast Advanced Master Mix (Life Technologies, Foster City, CA, USA) and PrimeTime^{TM} assay (Integrated DNA Technologies, Coralville, IA, USA), in a final reaction volume of 10 µL. Standard cycling conditions were used to analyze samples in a 7900HT Fast Real-Time PCR system (Life Technologies, Foster City, CA, USA). Expression of 8 candidate biomarker genes for each gender was studied and Cq values were collected with SDS 2.3 software. Minimum

Information on qPCR Experiments (MIQE) guidelines were taken into account. (Bustin et al. 2009) Amplification efficiencies were between 90-110% for all assays. Raw data were processed to normalized relative gene expression values with qBase plus (Biogazelle, Zwijnaarde, Belgium) (Hellemans et al. 2007). Triplicates were run for all samples; technical replicates were included when the difference in Cq value was < 0.5. A set of three genes was used for data normalization, namely Hypoxanthine Phosphoribosyltransferase 1 (*HPRT*), Importine 8 (*IPO8*) and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta (*YWHAZ*).

Data analysis

Statistical analyses were carried out using SAS software (version 9.3, SAS Institute Inc., Cary, NC, USA). Continuous data were presented as mean ± standard deviation (SD) and categorical data as percentages (%) and frequencies. Models were adjusted for age, body mass index (BMI), SES, smoking (categorized as smokers, former smokers and never smokers), white blood cell counts (absolute number of leukocytes and percentage of neutrophils), time of day (<12pm, 12–3pm, 3–6pm, >8pm) and season (October–March or April–September) of blood sampling. P-values <0.05 were considered statistically significant, p-values corrected for multiple testing referred to as q-values. We plotted residuals for each gene to check whether significance was driven by outliers, these were removed were appropriate. To indicate significance of selected biomarker genes for each gender, we included an interaction term for gender in our main analysis. P-values for the interaction term gender were calculated for all genes under study, not only those that were significant.

In validation analysis, we examined the association between gene expression and PM_{10} exposure, stratified by sex using linear regression models for the 8 selected genes for each gender.

ROC Curves exposure prediction

We calculated the ability to predict PM_{10} exposure based on expression of the set of 8 validated genes significantly associated with PM_{10} exposure in the discovery cohort for each gender. For this purpose, we estimated sensitivity and specificity of the prediction using receiver operating characteristic (ROC) plots. Subjects were stratified according to their long-term PM_{10} exposure levels with the 75th percentile as cut-off point (25.7 $\mu g/m^3$ annual mean for women, 24.5 $\mu g/m^3$ for men). All analyses were repeated similarly using long-term $PM_{2.5}$ exposure levels, the cut-off point for long-term PM2.5 exposure, or the 75th percentile of exposure was 16.0 $\mu g/m^3$ for both men and women.

Results

Population characteristics

Table 1 lists the characteristics of the study cohorts. All participants were of European origin. Distribution of sex, SES, age and BMI as well as exposure did not differ between the discovery and validation cohort. Both cohorts included just less than 50% men and age averaged (SD) 57.9 (4.3) years. Season of sampling differed between both cohorts, with sampling for the discovery phase of the study mainly occurring throughout the warm months of the year, whereas sampling for the validation study was mainly performed during the cold season. However, since we are working with average annual exposures over a 2-year period, this approach in itself corrects for the differences across seasons. Blood sampling was done ≤3 pm for all discovery cohort participants, while most validation cohort participants had samples drawn after 3pm. The discovery cohort consistent only of non-smokers, whereas the validation cohort included smokers (n=21).

Gene level analysis

Table 2 displays the 20 top genes for PM_{10} and $PM_{2.5}$ exposure in men and women. Excel File Tables S1–S4 display the extended top 50 lists for each exposure/gender combination. An overview on the total number of significant genes identified in our analysis, indicating the overlap between men and women, is given in Figure 2. For 199 gene transcripts we noticed significant sex by particulate matter exposure (PM_{10}) interactions (data not shown). The corresponding number of gene transcripts for $PM_{2.5}$ with a significant sex by exposure interaction was 601 (data not shown). In men, there were significant associations between 47 genes and PM_{10} only, 149 genes and $PM_{2.5}$ only, and there were 92 genes associated with both exposures. In women there were significant associations between 91 genes and PM_{10} only, 1067 genes and $PM_{2.5}$ only, and there were 498 genes associated with both exposures. We identified two genes in common between long-term PM_{10} exposure in men and women, namely *RAC3* and *DNAJB5*, respectively ranked as the 290th and 331th gene most significant genes with PM_{10} exposure in women (out of 592 genes). Furthermore *RAC3* was also significantly associated with long-term $PM_{2.5}$ exposure in men and *DNAJB5* with long-term PM2.5 exposure in women. We did not observe any significant FDR-corrected q-values in the discovery phase of our study.

Pathway analysis

There were 1,251 and 966 pathways significantly associated with PM10 and PM2.5, respectively, in men, and 280 and 182 pathways significantly associated with PM10 and PM2.5 in women, based on uncorrected p-values. The top 5 identified pathways for each indicator of exposure are summarized in Table 3.

Long-term PM_{10} exposure in men is associated with response to elevated platelet cytosolic Ca^{2+}, the prolactin signaling pathway and platelet degranulation. The 5th top significant pathway in association with PM_{10} exposure in men is signaling by insulin receptor, which

ranks 4th when analyzing long-term $PM_{2.5}$ exposure. Other pathways associated with $PM_{2.5}$ exposure in men are cell-cell communication and signaling by Type 1 Insulin-like Growth Factor and Insulin receptor signaling cascade. For women, long-term PM_{10} exposure was associated with, in descending order of significance, respiratory electron transport, packaging of telomere ends, electron transport chain, respiratory electron transport and telomere maintenance. $PM_{2.5}$ exposure was associated with respiratory electron transport, and the proteasome in women (Table 3).

Transcriptome signature in relation to long-term exposure

We selected 8 genes that were significantly ($p<0.05$) associated with long-term PM_{10} exposure in the microarray study and have a published link with air pollution-related disease (Table 4) for validation in an independent cohort. Of these we could confirm (i.e. they were also significantly associated with PM_{10} in the validation cohort based on uncorrected p-values, and associations were in the same direction as in the discovery cohort) 2 out of 8 genes for men [DnaJ homolog, subfamily B, member 5 (*DNAJB5*), and E2F associated phosphoprotein (*EAPP*)] and 1 out of 8 genes for women to be [Rho GTPase Activating protein 4 (*ARHGAP4*) borderline significantly ($p=0.0535$) associated with PM_{10} exposure (Table 4). *AKAP6* ($p = 0.02$) and *LIMK1* ($p = 0.006$) were significantly associated with PM10 in women in the validation cohort, albeit with significantly lower expression instead of higher expression as in the discovery cohort. We also tested the same sets of 8 genes for each sex for associations with $PM_{2\cdot5}$ exposure in the validation cohort, since all but one of the candidate genes (*PYG02* in women, which also was not significant for PM10 in the discovery cohort) were significantly associated with long-term $PM_{2\cdot5}$ exposure in the discovery cohort. For $PM_{2.5}$ exposure, we could confirm 2 out of 8 genes [*DNAJB5* (borderline significant, $p=0.059$)

and *EAPP*] for men and 4 out of 8 genes for women [*ARHGAP4*, *PYGO2*, sirtuin 7 (*SIRT7*) and Autophagy related 16-like 2 (*ATG16L2*)] (see Supplemental Material, Table S1). Excluding 21 current smokers (14 of 94 women and 7 of 75 men) from the validation cohort did not alter our conclusions, based on the similarity in the effect estimates, apart for expression of *ARHGAP4* in association with long-term PM_{10} exposure (Supplemental Material, Table S2).

Validation set

To determine whether gene expression candidate biomarkers identified in the discovery cohort were robust exposure markers, we performed ROC curve analysis with long-term PM_{10} exposure level 24.5 $\mu g/m^3$ (75^{th} percentile) as cut-off point in men. Figure 3. A shows the sensitivity and 1 minus specificity (false positive ratio) of PM_{10} exposure levels for men in association with the candidate biomarker genes. The model including the 8 genes in men had an area under the curve (AUC) value of 0.92 (95% CI: 0.85, 1.00; p=0.0002) . In women the model including the 8 genes had an AUC of 0.86 (95% CI: 0.76, 0.96; p=0.004) (Figure 3, panel B, cut-off point 25.7 $\mu g/m^3$). The combined geneset perfomed better both in men and women than the individual genes. Similarly, for $PM_{2.5}$ exposure prediction, the model for men had an AUC of 0.91 (95% CI: 0.83, 0.97; p=0.007) (Figure 3, panel C), the model for women had an AUC of 0.90 (95 % CI: 0.81, 0.98; p=0.0002) (Figure 3, panel D).

Discussion

We identified and validated transcriptome signatures that are associated with long-term exposure to particulate air pollution in apparently healthy men and women. These sets of 8 gender-specific genes were predictive of exposure in the validation cohort, and including all 8 genes in one model provided a better prediction than the 8 genes individually. We found *DNAJB5* and *EAPP* in men and *ARHGAP4* in women based on a discovery set and a

validation analysis to be significantly associated with PM_{10} exposure. When analyzing $PM_{2.5}$ exposure, for women, besides *ARHGAP4* we identified *PYGO2*, *SIRT7* and *ATG16L2* as significantly associated with particulate matter exposure. However, we cannot assume these associations indicate causal relations due to the observational nature of our study. ROC analysis revealed excellent separation between individuals with high and low exposure to long-term particulate air pollution using the genes selected for validation. We believe gene expression levels have potential to be used as biomarker of exposure and effect with high specificity to link particulate air pollution to its health consequences, as these can be measured at the personal level rather than be obtained through exposure modelling at the population level. Further studies looking at different age and ethnic groups are warranted to explore the capabilities of gene expression levels as predictors in more depth. Longitudinal studies that monitor disease incidence, exposure and gene expression over time would be excellent to provide more insights.

We observed different transcriptomic expression levels in association with particulate air pollution exposure in men and women. Sex-specific differences may be explained by differences in inflammatory responses between men and women. Immunologic differences between men and women have been reported based on gene expression profiles in blood between smokers and non-smokers, where women seem to have a more specific (involving less extensive pathways) immunologic response to smoking than men (Faner et al. 2014). Furthermore, sex-specific associations were also reported for microarray expression profiles in relation to environmental exposure to diverse compounds such as polychlorinated biphenyls, dioxin, benzene and PAHs (De Coster et al. 2013). The sex-specific associations between PM and gene expression that we observed are in line with previous reports of sex-specific associations with other exposures. As such, prenatal exposure to bisphenol A (BPA) led to differential responses in murine placentae of female and male embryos. (Imanishi et al.

2003) Prenatal stress exposure in rats was associated with sex-specific differences in gene expression and behavioral effects in male and female offspring (Van den Hove et al. 2013). This study clearly shows the same biological exposure (i.e. prenatal stress) leads to a highly differential response in male and female offspring.

To date, limited human data is available on microarray gene expression profiling in response to air pollution exposure. However, in an attempt to study the effects of *in utero* carcinogenic exposures, gene expression profiles in cord blood from 111 babies participating in the Norwegian BraMat cohort were assessed and correlation analyses of gene expression levels with biomarkers of exposure measured showed variable numbers of significantly correlating genes. Overall, separate analyses for male and female newborns resulted in higher numbers of significantly correlating genes per gender with low overlap of similarly expressed genes between the 2 sexes, thus indicating a clear gender-specific toxicogenomic response. More specifically, the authors reported only 1 gene in common between girls (39 significant genes) and boys (331 significant genes) for dioxin exposure (Hochstenbach et al. 2012).

Given evidence of the differential responses to PM exposure both at the gene and pathway levels between men and women, we hypothesize that different pathways could lead to the same disease outcome in both genders. Recently, it was reported that the same personal exposure (i.e. smoking) could lead to disease in a differential manner in men and women. As such, Paul et al. (2014) described microarray analysis in smokers and non-smoking men and women. They utilized a population of 24 middle-aged smoking men (n=12) and women (n=12) and an equal number of non-smoking controls. The gene set correlated with smoking in men was incapable of separating female smokers from non-smokers and vice versa. They identified a large number of oncogenic pathway gene-sets that were significantly different in female smokers compared to male smokers with Gene Set Enrichment Analysis of microarray data. In addition, functional annotation with Ingenuity Pathway Analysis (IPA) identified

smoking-correlated genes associated with biological functions in male and female smokers that are directly relevant to well-known smoking related pathologies. However, these relevant biological functions were overrepresented in female smokers compared to male smokers. Identified pathway categories in women were xenobiotic metabolism signaling, actin metabolism signaling, clathrin-mediated signaling, eicosanoid signaling, thrombin signaling, tight junction signaling, molecular mechanism of cancer and natural killer cell signaling (Paul and Amundson 2014).

The expression of *ARHGAP4* was borderline significantly associated with long-term PM_{10} exposure in women in the discovery cohort, and borderline significant (p=0.0535) in the validation cohort. *ARHGAP4, SIRT7* and *ATG16L2* were furthermore significantly associated with long-term $PM_{2.5}$ exposure in women in the discovery cohort and validation cohort.

ARHGAP4 is a RhoGAP that regulates the cytoskeletal dynamics that controls cell motility and axon outgrowth (Vogt et al. 2007). Pygosus 2 (*PYGO2*) is a component of the Wnt signaling pathway required for β-catenin/T-cell factor (TCF)-dependent transcription and has been shown to be upregulated in lung cancer both *in vitro* in non-small cell lung cancer cell lines and *in vivo* in human primary tumor tissue samples (Zhou et al. 2014).

In vitro experiments using hematopoietic stem cells from sirtuin 7 (*SIRT7*) knockout mice have shown *SIRT7* regulates mitochondrial activity and its inactivation causes reduced quiescence, increased mitochondrial protein folding stress, and compromised regenerative capacity of hematopoietic stem cells (Mohrin et al. 2015),(Liu and Chen 2015). Mitochondrial DNA and function have been shown to be associated with chronic air pollution exposure in populations of newborns (Janssen et al. 2012) and elderly men (Zhong et al. 2016), hence NAD-dependent deacetylase SIRT7 might provide insight into a molecular mechanism underlying the mitochondrial damage following air pollution exposure. Autophagy related 16-like 2 (*ATG16L2*) is a core autophagy gene. Previously, we found in newborns epigenetic

modifications in the mitochondrial genome, in association with $PM_{2.5}$ exposure during gestation and placental mtDNA content, which could reflect signs of mitophagy and mitochondrial death (Janssen et al. 2012).

The expression of the genes *DNAJB5* and *EAPP* were significantly associated with PM_{10} air pollution exposure in men, in the discovery cohort and validation cohort. DNAJB5 is a member of the evolutionarily conserved DNAJ/HSP40 family of proteins, which regulate molecular chaperone activity by stimulating ATPase activity (Ohtsuka and Hata 2000). DNAJB5 contains a cysteine-rich domain which renders the protein sensitive to ROS. The protein forms a multiprotein complex together with Trx1 and class II histone deacetylases (HDACs) that functions as a master negative regulator of cardiac hypertrophy (Ago et al. 2008). E2F-associated phospho-protein (EAPP) is a nuclear phosphoprotein that interacts with the activating members of the E2F transcription factor family. *In vitro* overexpression of EAPP increased the fraction of G1 cells and led to heightened resistance against DNA damage. EAPP itself becomes upregulated after DNA damage and stimulates the expression of p21 independently of p53 (Andorfer and Rotheneder 2011).

In pathway analyses, we identified several respiratory chain related pathways significantly associated with long-term PM_{10} and $PM_{2.5}$ exposure in women. Rossner *et al.* (Rossner et al. 2015) reported deregulation of expression of respiratory chain, oxidative phosphorylation and mitochondrial membrane pathways when comparing gene expression profiles in adult non-smoking men from a heavily polluted area versus a control region in the Czech Republic across different seasons (winter and summer 209 and winter 2010).

Although sex-related differences have been observed for different environmental pollutions, to our knowledge, this is the first study on microarray gene expression profiles in association with long-term air pollution exposure among middle aged men and women.

Our study has strengths and limitations. We did our investigations in two independent cohorts for discovery and validation, using the same exposure modeling and used the gold standard qPCR as validation tool (Canales et al. 2006). Although sample size for the discovery cohort was limited, we believe validation in an independent cohort based on a reliable method such as qPCR indicates the robustness of our analyses. Our study also has its limitations inherent to the cross-sectional nature of our study. We used 2010–12 air pollution data to develop our high-resolution exposure models, which we applied to the participants' baseline addresses (2004). Studies in the Netherlands (Brauer et al. 2003), Italy (Rome) (Rosenlund et al. 2008), the UK (Briggs et al. 2000), and Canada (Vancouver) (Henderson et al. 2007) have shown that during periods of about 10 years and longer, existing land use regression models predicted historic spatial contrasts well. The use of a relatively homogenous population limits the potential generalizability of our study to populations with different ages, races/ethnicities or locations. Lastly, our study design did not allow to control for cell counts in the discovery phase of the study. As cell counts were not performed on the samples for microarray analysis, and there is no good means for imputation of these values for Agilent 4X44K arrays during data analysis, we were not able to control for this.

This is the first time levels of gene expression of candidate genes have been used to accurately predict air pollution exposure levels ($PM_{10,}$ $PM_{2.5}$). For this purpose, we have established ROC curves based on the genes selected for validation in an independent cohort, and were able to separate low (<75th percentile) from high (>75th percentile) exposed individuals. ROC curves are commonly used to compare the diagnostic performance of two or more tests, as they give a good indication of both the sensitivity and specificity of the studied test (Greiner et al. 2000). As such, it has been demonstrated that gene expression signatures can predict survival for instance in pancreatic (Newhook et al. 2014) or non-small cell lung cancer (Lu et al. 2006). In 2009, this technique was applied for the first time in an

environmental epidemiology setting, showing that specific DNA methylation patterns could accurately predict the relationship between exposure to airborne PAHs and childhood asthma incidence. Perera et al (2009) investigated PAH levels in cord blood samples from 20 newborns and replicated the association between PAH levels and candidate region methylation in 56 other newborns from the CCCEH cohort that recruits nonsmoking Dominican and African American women and their children residing in different areas of New York, USA (Perera et al. 2009). However, the application of this approach to the field of gene expression data in association with air pollution exposure is novel.

In ROC curve analysis, an AUC of 0.80 is considered a ROC curve with good separation characteristics, and an AUC of 0.90 is considered excellent, in its ability to distinguish between true- and false positives. We have identified sex-specific gene-sets that fulfill these criteria for PM_{10} and $PM_{2.5}$ exposure. However, we must interpret the current set within the context of its limitations inherent to its cross-sectional nature of our study.

Conclusions

In conclusion, microarray analysis has identified different gene expression levels in response to long-term air pollution in men and women. From gene-level analysis, candidate biomarker genes with a reported link to AP-related disease were selected and validated (i.e. significantly associated with PM exposure with the same direction of regulation of expression) in an independent cohort. For men, we propose *DNAJB5* and *EAPP* as biomarkers of exposure. For women we identified *ARHGAP4*, *PYGO2*, *SIRT7* and *ATG16L2* as biomarker genes of exposure. ROC analysis revealed that the genes were able to predict high or low PM_{10} exposure accurately. Prospective studies in other populations are needed to confirm our findings with regard to sex-specific expression of these genes in association with PM exposure. Furthermore, it would be highly relevant to analyze the gene expression of these

gender-specific gene-sets in cohorts with higher PM exposure as well as in subjects at different stages of life, including the more vulnerable stages such as early childhood and puberty.

References

Ago T, Liu T, Zhai P, Chen W, Li H, Molkentin JD, et al. 2008. A redox-dependent pathway for regulating class II HDACs and cardiac hypertrophy. Cell 133(6): 978-993.

Andorfer P, Rotheneder H. 2011. EAPP: gatekeeper at the crossroad of apoptosis and p21-mediated cell-cycle arrest. Oncogene 30(23): 2679-2690.

Alfaro-Moreno E, Nawrot TS, Nemmar A, Nemery B. 2007. Particulate matter in the environment: pulmonary and cardiovascular effects. Curr Opin Pulm Med 13(2): 98-106.

Bowatte G, Lodge C, Lowe AJ, Erbas B, Perret J, Abramson MJ, et al. 2015. The influence of childhood traffic-related air pollution exposure on asthma, allergy and sensitization: a systematic review and a meta-analysis of birth cohort studies. Allergy 70(3): 245-256.

Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, et al. 2003. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. Epidemiology. 14: 228–239.

Briggs DJ, de Hoogh C, Gulliver J, Wills J, Elliott P, Kingham S, et al. 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. Sci. Total. Environ. 253 (1–3), 151–167.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55(4): 611-622.

Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C, et al. 2006. Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotechnol 24(9): 1115-1122.

Chen Q, Jiao D, Hu H, Song J, Yan J, Wu L. et al. 2013. Downregulation of LIMK1 level inhibits migration of lung cancer cells and enhances sensitivity to chemotherapy drugs. *Oncology research* 20(11):491-498.

Chen JC, Schwartz J. 2009. Neurobehavioral effects of ambient air pollution on cognitive performance in US adults. Neurotoxicology 30(2): 231-239.

Cosemans, G., Kretzschmar, J.G., De Baere G. and Vandervee J. 1982. Large Scale Validation of a bi-Gaussian Dispersion Model in a Multiple Source Urban and Industrial Area, in De Wispelaere C. (Editor), Air Pollution Modeling and Its Application II, Plenum Press; pp. 709-728.

Cosemans, G., Kretzschmar, J., Janssen, L. and Maes G. 1995. The Third Workshops environmental impact assessment model intercomparison exercise, International Journal Environment and Pollution 5(4-6): 785-798.

Cosemans, G., Ruts R. and Kretzschmar J.G. 2001. Impact assessment with the Belgian dispersion model IFDM and the New Dutch National Model, Proc. 7th int. conf. on Harmonisation within atmospheric dispersion modelling for regulatory purposes, Belgirate, Italy, 125-129.

Cosin-Sales J, Christiansen M, Kaminski P, Oxvig C, Overgaard MT, Cole D, et al. 2004. Pregnancy-associated plasma protein A and its endogenous inhibitor, the proform of eosinophil major basic protein (proMBP), are related to complex stenosis morphology in patients with stable angina pectoris. *Circulation* 109(14):1724-1728.

Dadvand P, Nieuwenhuijsen MJ, Esnaola M, Forns J, Basagana X, Alvarez-Pedrerol M, et al. 2015. Green spaces and cognitive development in primary schoolchildren. Proc Natl Acad Sci U S A 112(26): 7937-7942.

De Boever P, Louwies T, Provost E, Int Panis L, Nawrot TS. 2014. Fundus photography as a convenient tool to study microvascular responses to cardiovascular disease risk factors in epidemiological studies. Journal of visualized experiments : JoVE(92): e51904.

De Coster S, van Leeuwen DM, Jennen DG, Koppen G, Den Hond E, Nelen V, et al. 2013. Gender-specific transcriptomic response to environmental exposure in Flemish adults. Environ Mol Mutagen 54(7): 574-588.

de Curtis I. 2008. Functions of Rac GTPases during neuronal development. Dev Neurosci 30(1-3): 47-58.

DeMuth JP, Jackson CM, Weaver DA, Crawford EL, Durzinsky DS, Durham SJ et al. 1998. The gene expression index c-myc x E2F-1/p21 is highly predictive of malignant phenotype in human bronchial epithelial cells. American journal of respiratory cell and molecular biology 19(1):18-24.

Doherty CJ, Kay SA. 2010. Circadian control of global gene expression patterns. Annu Rev Genet 44: 419-444.

Donaldson K, Tran L, Jimenez LA, Duffin R, Newby DE, Mills N, et al. 2005. Combustion-derived nanoparticles: a review of their toxicology following inhalation exposure. Part Fibre Toxicol 2: 10.

Faner R, Gonzalez N, Cruz T, Kalko SG, Agusti A. 2014. Systemic inflammatory response to smoking in chronic obstructive pulmonary disease: evidence of a gender effect. PLoS One 9(5): e97491.

Fink SP, Yamauchi M, Nishihara R, Jung S, Kuchiba A, Wu K, et al. 2014. Aspirin and the risk of colorectal cancer in relation to the expression of 15-hydroxyprostaglandin dehydrogenase (HPGD). Sci Transl Med 6(233): 233re232.

Finkelstein MM, Jerrett M. 2007. A study of the relationships between Parkinson's disease and markers of traffic-derived and environmental manganese air pollution in two Canadian cities. Environ Res 104(3): 420-432.

Fujii T, Hayashi S, Hogg JC, Mukae H, Suwa T, Goto Y, et al. 2002. Interaction of alveolar macrophages and airway epithelial cells following exposure to particulate matter produces mediators that stimulate the bone marrow. Am J Respir Cell Mol Biol 27(1): 34-41.

Greiner M, Pfeiffer D, Smith RD. 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 45(1-2): 23-41.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol 8(2): R19.

Henderson, S., Beckerman, B., Jerrett, M., Brauer, M. 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environ. Sci. Technol. 41, 2422–2428.

Heredia L, Helguera P, de Olmos S, Kedikian G, Solá Vigo F, LaFerla F, et al. (2006) Phosphorylation of actin-depolymerizing factor/cofilin by LIM-kinase mediates amyloid beta-induced degeneration: a potential mechanism of neuronal dystrophy in Alzheimer's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 26(24):6533-6542.

Hochstenbach K, van Leeuwen DM, Gmuender H, Gottschalk RW, Lovik M, Granum B, et al. 2012. Global gene expression analysis in cord blood reveals gender-specific differences in response to carcinogenic exposure in utero. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer

Research, cosponsored by the American Society of Preventive Oncology 21(10): 1756-1767.

Huang YC, Schmitt M, Yang Z, Que LG, Stewart JC, Frampton MW, et al. 2010. Gene expression profile in circulating mononuclear cells after exposure to ultrafine carbon particles. Inhal Toxicol 22(10): 835-846.

Huang L, Poke G, Gecz J, Gibson K, 2012. A novel contiguous gene deletion of AVPR2 and ARHGAP4 genes in male dizygotic twins with nephrogenic diabetes insipidus and intellectual disability. *American journal of medical genetics. Part A* 158A(10):2511-2518.

Husten L. 1998. More data reported for HDL's role in heart disease. *Lancet* 352(9140):1603.

Imanishi S, Manabe N, Nishizawa H, Morita M, Sugimoto M, Iwahori M, et al. 2003. Effects of oral exposure of bisphenol A on mRNA expression of nuclear receptors in murine placentae assessed by DNA microarray. The Journal of reproduction and development 49(4): 329-336.

Iwabayashi M, Taniyama Y, Sanada F, Azuma J, Iekushi K, Kusunoki H et al. 2012. Role of serotonin in angiogenesis: induction of angiogenesis by sarpogrelate via endothelial 5-HT1B/Akt/eNOS pathway in diabetic mice. Atherosclerosis 220(2):337-342.

Jacobs L, Emmerechts J, Mathieu C, Hoylaerts MF, Fierens F, Hoet PH, et al. 2010. Air pollution related prothrombotic changes in persons with diabetes. Environ Health Perspect 118(2): 191-196.

Jang I, Jeon BT, Jeong EA, Kim EJ, Kang D, Lee JS, et al. 2012. Pak1/LIMK1/Cofilin Pathway Contributes to Tumor Migration and Invasion in Human Non-Small Cell Lung Carcinomas and Cell Lines. Korean J Physiol Pharmacol 16(3): 159-165.

Janssen BG, Munters E, Pieters N, Smeets K, Cox B, Cuypers A, et al. 2012. Placental Mitochondrial DNA Content and Particulate Air Pollution during in Utero Life. Environ Health Perspect 120(9): 1346-1352.

Janssen S, Dumont G, Fierens F, Mensink C. 2008. Spatial interpolation of air pollution measurements using CORINE land cover data. Atmospheric Environment 42(20): 4884-4903.

Johnson JY, Rowe BH, Villeneuve PJ. 2010. Ecological analysis of long-term exposure to ambient air pollution and the incidence of stroke in Edmonton, Alberta, Canada. Stroke 41(7): 1319-1325.

Kalkstein LS, Valimont KM. 1986. An Evaluation of Summer Discomfort in the United-States Using a Relative Climatological Index. Bulletin of the American Meteorological Society 67(7): 842-848.

Kamburov A, Stelzl U, Lehrach H, Herwig R. 2012. The ConsensusPathDB interaction database: 2013 update. Nucleic Acids Research.

Kleensang A, Maertens A, Rosenberg M, Fitzpatrick S, Lamb J, Auerbach S, et al. 2014. t4 workshop report: Pathways of Toxicity. ALTEX 31(1): 53-61.

Ko FW, Tam W, Wong TW, Chan DP, Tung AH, Lai CK, et al. 2007. Temporal relationship between air pollutants and hospital admissions for chronic obstructive pulmonary disease in Hong Kong. Thorax 62(9): 780-785.

La Rocca C, Tait S, Guerranti C, Busani L, Ciardo F, Bergamasco B, et al. 2014. Exposure to endocrine disrupters and nuclear receptor gene expression in infertile and fertile women from different Italian areas. Int J Environ Res Public Health 11(10): 10146-10164.

Lefebvre W, Van Poppel M, Maiheu B, Janssen S, Dons EJS. 2013. Evaluation of the RIO-IFDM-street canyon model chain. Atmospheric Environment 77: 325-337.

Li SH, Luo YL and Lai WY. 2006. Expression of eosinophil major basic protein mRNA in bronchial asthma]. *Journal of Southern Medical University* 26(9):1330-1333.

Lin YC, Koleske AJ. 2010. Mechanisms of synapse and dendrite maintenance and their disruption in psychiatric and neurodegenerative disorders. Annu Rev Neurosci 33: 349-378.

Liu JP, Chen R. 2015a. Stressed SIRT7: facing a crossroad of senescence and immortality. Clin Exp Pharmacol Physiol 42(6): 567-569.

Liu Y, Dong QZ, Wang S, Fang CQ, Miao Y, Wang L et al. 2013. Abnormal expression of Pygopus 2 correlates with a malignant phenotype in human lung cancer. BMC cancer 13:346.

Liu TQ, Wang GB, Li ZJ, Tong XD, & Liu HX, 2015b. Silencing of rac3 inhibits proliferation and induces apoptosis of human lung cancer cells. Asian Pacific journal of cancer prevention :APJCP 16(7):3061-3065.

Ljubimova JY, Kleinman MT, Karabalin NM, Inoue S, Konda B, Gangalum P, et al. 2013. Gene expression changes in rat brain after short and long exposures to particulate matter in Los Angeles basin air: Comparison with human brain tumors. Exp Toxicol Pathol 65(7-8): 1063-1071

Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, et al. 2006. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. PLoS Med 3(12): e467.

Maes, G., Cosemans, G., Kretzschmar, J., Janssen L. and Van Tongerloo, J. 1995. Comparison of six Gaussian dispersion models used for regulatory purposes in different countries of the EU, IJEP, 5(4-6): 734-747.

Magne J, Gustafsson P, Jin H, Maegdefessel L, Hultenby K, Wernerson A et al. (2015) ATG16L1 Expression in Carotid Atherosclerotic Plaques Is Associated With Plaque Vulnerability. Arteriosclerosis, thrombosis, and vascular biology 35(5):1226-1235.

Magnussen H, Jorres R, Nowak D. 1993. Effect of air pollution on the prevalence of asthma and allergy: lessons from the German reunification. Thorax 48(9): 879-881.

Maiheu B, Veldeman B, Viaene P, De Ridde rK, Lauwaet D, Smeets N, et al. 2013. Identifying the best available large-scale concentration maps for air quality in Belgium.

[http://www.milieurapport.be/Upload/main/0_onderzoeksrapporten/2013/Eindrapport_Concentratiekaarten_29_01_2013_TW.pdf].

Melchior JT, Sawyer JK, Kelley KL, Shah R, Wilson RD, Hantgan RR et al. 2013. LDL particle core enrichment in cholesteryl oleate increases proteoglycan binding and promotes atherosclerosis. *Journal of lipid research* 54(9):2495-2503.

Mensink, C. and Maes, G. 1996. Comparative sensitivity study for operational short-range atmospheric dispersion models, IJEP 8(3-6): 356-366.

Mills NL, Donaldson K, Hadoke PW, Boon NA, MacNee W, Cassee FR, et al. 2009. Adverse cardiovascular effects of air pollution. Nat Clin Pract Cardiovasc Med 6(1): 36-44.

Mohrin M, Shin J, Liu Y, Brown K, Luo H, Xi Y, et al. 2015. Stem cell aging. A mitochondrial UPR-mediated metabolic checkpoint regulates hematopoietic stem cell aging. Science 347(6228): 1374-1377.

Montaner D, Tarraga J, Huerta-Cepas J, Burguet J, Vaquerizas JM, Conde L, et al. 2006. Next station in microarray data analysis: GEPAS. Nucleic Acids Res 34(Web Server issue): W486-491.

Nel AE, Diaz-Sanchez D, Li N. 2001. The role of particulate pollutants in pulmonary inflammation and asthma: evidence for the involvement of organic chemicals and oxidative stress. Curr Opin Pulm Med 7(1): 20-26.

Newhook TE, Blais EM, Lindberg JM, Adair SJ, Xin W, Lee JK, et al. 2014. A thirteen-gene expression signature predicts survival of patients with pancreatic cancer and identifies new genes of interest. PLoS One 9(9): e105631.

Ohtsuka K, Hata M. 2000. Mammalian HSP40/DNAJ homologs: cloning of novel cDNAs and a proposal for their classification and nomenclature. Cell Stress Chaperones 5(2): 98-112.

Oldenburger A, Poppinga WJ, Kos F, de Bruin HG, Rijks WF, Heijink IH, et al. 2014. A-kinase anchoring proteins contribute to loss of E-cadherin and bronchial epithelial barrier by cigarette smoke. Am J Physiol Cell Physiol 306(6): C585-597.

Olesen H. 1995. The Model Validation Exercise at Mol: Overview of Results, Int J of Environment and Pollution, 5 (4-6): 761-784.

Oti M, Snel B, Huynen MA, Brunner HG. 2006. Predicting disease genes using protein-protein interactions. Journal of medical genetics 43(8):691-698.

Paul S, Amundson SA. 2014. Differential Effect of Active Smoking on Gene Expression in Male and Female Smokers. Journal of carcinogenesis & mutagenesis 5.

Perera F, Tang WY, Herbstman J, Tang D, Levin L, Miller R, et al. 2009. Relation of DNA methylation of 5'-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. PloS one 4(2): e4488.

Pettit AP, Brooks A, Laumbach R, Fiedler N, Wang Q, Strickland PO, et al. 2012. Alteration of peripheral blood monocyte gene expression in humans following diesel exhaust inhalation. Inhal Toxicol 24(3): 172-181.

Pope CA, 3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. 2004. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 109(1): 71-77.

Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro A, 3rd, Schwartz J. 2011. Traffic-related air pollution and cognitive function in a cohort of older men. Environ Health Perspect 119(5): 682-687.

Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, Jensen SS, Ketzel M, Sorensen M, et al. 2011. Lung cancer incidence and long-term exposure to air pollution from traffic. Environ Health Perspect 119(6): 860-865.

Raz L, Zhang QG, Zhou CF, Han D, Gulati P, Yang LC, et al. 2010. Role of Rac1 GTPase in NADPH oxidase activation and cognitive impairment following cerebral ischemia in the rat. PLoS One 5(9): e12606.

Roepstorff K, Rasmussen I, Sawada M, Cudre-Maroux C, Salmon P, Bokoch G, et al. 2008. Stimulus-dependent regulation of the phagocyte NADPH oxidase by a VAV1, Rac1, and PAK1 signaling axis. J Biol Chem 283(12): 7983-7993.

Rosenlund M, Forastier F, Stafoggia M, Porta D, Perucci M, Ranzi A, Nussio F, Perucci CA. 2008. Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome. J. Expo. Sci. Environ. Epidemiol. 18, 192–199.

Rossner P, Jr., Tulupova E, Rossnerova A, Libalova H, Honkova K, Gmuender H, et al. 2015. Reduced gene expression levels after chronic exposure to high concentrations of air pollutants. Mutat Res 780: 60-70.

Rostila A, Puustinen A, Toljamo T, Vuopala K, Lindström I, Nyman TA et al. 2012. Peroxiredoxins and tropomyosins as plasma biomarkers for lung cancer and asbestos exposure. Lung cancer 77(2):450-459.

Scott RW, Olson MF. 2007. LIM kinases: function, regulation and association with human disease. J Mol Med (Berl) 85(6): 555-568.

Steadman RG. 1979. Assessment of Sultriness .1. Temperature-Humidity Index Based on Human Physiology and Clothing Science. Journal of Applied Meteorology 18(7): 861-873.

Vakhrusheva O, Smolka C, Gajawada P, Kostin S, Boettger T, Kubin T, et al. 2008. Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice. Circulation research 102(6):703-710.

van Breda SG, Wilms LC, Gaj S, Jennen DG, Briede JJ, Kleinjans JC, et al. 2015. The exposome concept in a human nutrigenomics study: evaluating the impact of exposure to a complex mixture of phytochemicals using transcriptomics signatures. Mutagenesis.

Van den Hove DL, Kenis G, Brass A, Opstelten R, Rutten BP, Bruschettini M, et al. 2013. Vulnerability versus resilience to prenatal stress in male and female rats; implications from gene expression profiles in the hippocampus and frontal cortex. European neuropsychopharmacology: the journal of the European College of Neuropsychopharmacology 23(10): 1226-1246.

van Leeuwen DM, Gottschalk RW, Schoeters G, van Larebeke NA, Nelen V, Baeyens WF, et al. 2008. Transcriptome analysis in peripheral blood of humans exposed to environmental carcinogens: a promising new biomarker in environmental health studies. EnvironHealth Perspect 116(11): 1519-1525.

Vineis P, van VK, Chadeau-Hyam M, Athersuch TJ. 2013. Advancing the application of omics-based biomarkers in environmental epidemiology. EnvironMolMutagen.

Vogt DL, Gray CD, Young WS, 3rd, Orellana SA, Malouf AT. 2007. ARHGAP4 is a novel RhoGAP that mediates inhibition of cell motility and axon outgrowth. Mol Cell Neurosci 36(3): 332-342.

Weuve J, Puett RC, Schwartz J, Yanosky JD, Laden F, Grodstein F. 2012. Exposure to particulate air pollution and cognitive decline in older women. Arch Intern Med 172(3): 219-227.

Young ME, Razeghi P, Taegtmeyer H. 2001. Clock genes in the heart: characterization and attenuation with hypertrophy. *Circulation research* 88(11):1142-1150.

Zhong J, Cayir A, Trevisi L, Sanchez-Guerra M, Lin X, Peng C, et al. 2016. Traffic-Related Air Pollution, Blood Pressure, and Adaptive Response of Mitochondrial Abundance. Circulation 133(4):378-87.

Zhou SY, Xu ML, Wang SQ, Zhang F, Wang L, Wang HQ. 2014. Overexpression of Pygopus-2 is required for canonical Wnt activation in human lung cancer. Oncology letters 7(1): 233-238.

Zhu X, Raina AK, Boux H, Simmons ZL, Takeda A, Smith MA. 2000. Activation of oncogenic pathways in degenerating neurons in Alzheimer disease. Int J Dev Neurosci 18(4-5): 433-437.

Table 1: Study population and exposure characteristics.

Characteristics	Discovery cohort	Validation cohort	Discovery cohort	Validation cohort
	Men (n= 48)	**Men (n= 75)**	**Women (n=50)**	**Women (n=94)**
Personal				
Age, Years	58.0±4.5	58.0±4.1	57.8±4.2	58.1±4.0
Body Mass Index, kg/m²	27.4±3.5	26.1±3.8	25.8±3.7	25.5±4.7
Socio-economic status				
Low	20 (41.7)	14 (18.7)	28 (56.0)	23 (24.5)
Medium	15 (31.3)	26 (34.7)	7 (14.0)	16 (17.0)
High	13 (27.1)	35 (46.7)	15 (30.0)	55 (58.5)
Smoking status				
Non-smokers	48 (100.0)	25 (33.3)	50 (100.0)	49 (52.1)
Former smoker	-	43 (57.3)	-	31 (33)
Current smoker	-	7 (9.3)	-	14 (14.9)
Season of blood sampling				
Cold (October-March)	40 (83.3)	27 (36.0)	40 (80.0)	40 (42.6))
Warm (April-September)	8 (16.7)	48 (64.0)	10 (20.0)	54 (57.4))
Time of blood sampling				
<12pm	41 (85.4)	0 (0.0)	44 (88.0)	7 (7.5)
12pm-3pm	7 (14.6)	20 (26.7)	6 (12.0)	25 (26.6)
3pm-6pm	0 (0.0)	32 (42.7)	0 (0.0)	43 (45.7)
>8pm	0 (0.0)	23 (30.7)	0 (0.0)	19 (20.2)
White blood cell count				
Leukocytes (#/μL)	-	6981.5±1632.1	-	6981.5±1632.1
Neutrophils (%)	-	56.8±8.1	-	56.8±8.1
Exposure (μg/m³)				
PM_{10} long-term	25.8 (21.5-30.4)	23.1 (20.3-27.4)	26.0 (20.5-35.3)	24.2 (20.4-28.2)
$PM_{2.5}$ long-term	17.7 (15.5-20.8)	15.5 (14.5-17.6)	17.8 (15.4-20.9)	16.0 (14.7-18.3)

Data are mean±SE or number (%), exposure data are mean (5-95th percentile)

Environ Health Perspect DOI: 10.1289/EHP370
Advance Publication: Not Copyedited

Table 2: Top 20 significant genes in association with 5 μg/m^3 increase in long-term PM_{10} and $PM_{2.5}$ exposure for men and women.

	Men				**Women**			
	PM_{10}		**$PM_{2.5}$**		**PM_{10}**		**$PM_{2.5}$**	
Rank no. gene	**Gene**	**FC (95% CI)**	**Gene**	**FC (95% CI)**	**Gene**	**FC (95% CI)**	**Gene**	**FC (95% CI)**
1	*EAPP*	1.15 (1.07 ,1.24)	*ISL2*	2.45 (1.58 ,3.78)	*ATG16L2*	0.81 (0.73 ,0.90)	*EFNB1*	0.64 (0.53 ,0.77)
2	*DCTN6*	1.23 (1.10 ,1.38)	*HDLBP*	1.31 (1.14 ,1.50)	*EFNB1*	0.79 (0.69 ,0.89)	*SLC6A7*	1.52 (1.25 ,1.86)
3	*DNAJB5*	1.36 (1.14 ,1.63)	*B3GNT3*	1.42 (1.18 ,1.70)	*SYTL1*	0.86 (0.79 ,0.93)	*FXN*	0.73 (0.62 ,0.85)
4	*ISL2*	1.55 (1.17 ,2.06)	*RNF144*	1.83 (1.28 ,2.62)	*SMG5*	0.84 (0.76 ,0.92)	*SFPQ*	1.41 (1.18 ,1.67)
5	*KIAA1914*	1.23 (1.07 ,1.42)	*ATOH8*	2.24 (1.37 ,3.66)	*TBC1D10C*	0.85 (0.78 ,0.93)	*NACAL*	0.69 (0.58 ,0.84)
6	*HDLBP*	1.14 (1.04 ,1.24)	*RAC3*	1.62 (1.21 ,2.18)	*NACAL*	0.81 (0.72 ,0.91)	*ATG16L2*	0.72 (0.61 ,0.86)
7	*B3GNT3*	1.19 (1.06 ,1.34)	*ADCK1*	1.49 (1.17 ,1.91)	*NFKBIE*	0.85 (0.78 ,0.93)	*SLC24A2*	1.73 (1.28 ,2.32)
8	*ATOH8*	1.55 (1.14 ,2.10)	*DNAJB5*	1.62 (1.20 ,2.18)	*CEMP1*	0.80 (0.70 ,0.91)	*THEX1*	0.34 (0.19 ,0.62)
9	*LSM12*	0.86 (0.77 ,0.95)	*ALX3*	1.40 (1.13 ,1.73)	*DCUN1D2*	0.84 (0.76 ,0.93)	*TBC1D13*	0.78 (0.68 ,0.90)
10	*ZNF187*	1.16 (1.04 ,1.28)	*MAN2A2*	1.39 (1.13 ,1.72)	*SLC6A7*	1.25 (1.10 ,1.43)	*VAPB*	1.21 (1.09 ,1.35)
11	*ARHGAP25*	1.11 (1.03 ,1.20)	*DCTN6*	1.35 (1.11 ,1.64)	*DHRSX*	1.21 (1.08 ,1.36)	*TPM3*	0.44 (0.28 ,0.70)
12	*SERF1B*	0.83 (0.72 ,0.95)	*DAK*	1.34 (1.10 ,1.64)	*TBC1D13*	0.86 (0.79 ,0.94)	*CYB5D1*	0.39 (0.23 ,0.67)
13	*ANXA1*	1.19 (1.05 ,1.36)	*PER1*	1.37 (1.11 ,1.69)	*SFPQ*	1.21 (1.08 ,1.35)	*ZNF77*	0.61 (0.46 ,0.81)
14	*TKTL1*	1.36 (1.09 ,1.71)	*GUCA2B*	1.78 (1.20 ,2.62)	*MAPK3*	1.19 (1.07 ,1.32)	*GABRD*	0.40 (0.24 ,0.67)
15	*PRG2*	1.29 (1.07 ,1.56)	*ATXN7L3*	1.30 (1.09 ,1.55)	*ZFYVE27*	0.91 (0.86 ,0.96)	*NFKBIE*	0.78 (0.68 ,0.90)
16	*PER1*	1.19 (1.05 ,1.36)	*LSM12*	0.77 (0.64 ,0.92)	*SLC39A2*	1.32 (1.11 ,1.55)	*CEACAM3*	1.67 (1.24 ,2.23)
17	*GUCA2B*	1.38 (1.09 ,1.75)	*PRG2*	1.57 (1.15 ,2.13)	*TSPAN4*	1.37 (1.13 ,1.65)	*TSPAN4*	1.69 (1.25 ,2.28)
18	*ST14*	1.20 (1.05 ,1.37)	*ABL2*	1.40 (1.11 ,1.78)	*DNAJC5*	0.87 (0.81 ,0.95)	*GPR137*	0.64 (0.50 ,0.83)
19	*CDV3*	0.84 (0.73 ,0.96)	*MAST3*	1.27 (1.07 ,1.49)	*MIA*	0.82 (0.73 ,0.93)	*DNAJC5*	0.80 (0.70 ,0.91)
20	*TTC30B*	1.20 (1.04 ,1.37)	*PIK3R1*	1.46 (1.12 ,1.89)	*CES2*	1.16 (1.06 ,1.28)	*HSF1*	0.85 (0.77 ,0.93)

FC= fold change. Rank no. gene indicates its hierarchy for that particular exposure and gender based on level of significance of the identified association, so gene ranked as no. 1 has the lowest P-value.

Table 3. The top 5 significant pathways defined by gene set enrichment analysis for each indicator of exposure.

Exposure	Pathway	Q-value	#measured/# genes in pathway
Men			
PM_{10}			
	Response to elevated platelet cytosolic Ca2+	3·11E-07	76/87
	Prolactin signaling pathway	5·78E-07	61/72
	Platelet degranulation	5·90E-07	71/82
	Leukocyte transendothelial migration	1·25E-06	98/118
	Signaling by Insulin receptor	5·18E-06	89/109
$PM_{2·5}$			
	Cell-Cell communication	1·35E-08	95/130
	Chagas disease (American trypanosomiasis)	1·40E-06	92/104
	Signaling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)	1·40E-06	76/96
	Signaling by Insulin receptor	1·93E-06	96/120
	Insulin receptor signaling cascade	2·33E-06	74/76
Women			
PM_{10}			
	Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins	2·08E-04	89/97
	Packaging Of Telomere Ends	3·98E-04	46/53
	Electron Transport Chain	8·11E-04	94/103
	Respiratory electron transport	9·59E-04	71/76
	Telomere Maintenance	1·50E-03	72/81
$PM_{2·5}$			
	Respiratory electron transport	9·07E-04	81/92
	Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins	1·77E-03	99/113
	Packaging Of Telomere Ends	4·54E-03	45/52
	Proteasome	4·93E-03	41/44
	Transcriptional regulation by small RNAs	4·93E-03	95/106

Pathways were identified using the Gene Set Enrichment Analysis Tool from the online Consensus Pathway Data Base.

Table 4. Selection of biomarker candidate genes, and their fold changes for an increase of 5 μg/m^3 long-term PM_{10} exposure.

Gene name	Gene description	Gene function	Link to disease	Discovery cohort FC (95% CI)	p-value	Validation cohort FC (95% CI)	p-value	q-value
Men								
DNAJB5	DnaJ (Hsp40) homolog, subfamily B, member 5	Heat shock protein 40	CVD (Ago, 2008)	1.36 (1.14, 1.63)	0.0014	1.64 (1.20, 2.23)	0.0026	0.02
RAC3	ras-related C3 botulinum toxin substrate 3 (rho family, small GTP binding protein Rac3)	Regulation of cellular responses (cell growth)	Lung cancer (Liu, 2015b)	1.25 (1.04, 1.51)	0.024	1.26 (0.94, 1.96)	0.10	0.18
EAPP	E2F associated phosphoprotein	Cell cycle/Apoptosis	Lung cancer (DeMuth, 1998)	1.15 (1.0, 1.24)	0.00055	1.18 (1.02, 1.38)	0.028	0.12
HDLBP	high density lipoprotein binding protein (vigilin)	Sterol metabolism	CVD (Husten, 1998)	1.14 (1.04, 1.24)	0.0065	1.02 (0.88, 1.19)	0.75	0.86
PRG2	Proteoglycan 2	Eosinophil major basic protein	CVD (Melchior, 2013), asthma (Li, 2006)	1.29 (1.07, 1.56)	0.012	1. 29 (0.98, 1.71)	0.066	0.18
PER1	period homolog 1 (Drosophila)	Circadian rhythm	CVD (Young, 2001)	1.19 (1.05, 1.36)	0.012	0.95 (0.74, 1.23)	0.72	0.86
PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (p85 alpha)	Insulin metabolism	Lung cancer (Lu, 2006)	1.22 (1.03, 1.43)	0.023	1.01 (0.82, 1.26)	0.91	0.91
SLA2	Src-like adaptor 2	SLAP adapter protein	CVD [73](Cosin-Sales, 2004)	1.22 (1.03, 1.44)	0.027	1.16 (0.97, 1.39)	0.11	0.18
Women								
AKAP6	A kinase (PRKA) anchor protein 6	Regulatory subunit of protein kinase A	CVD (Oti, 2006)	1.21 (1.07, 1.36)	0.0036	0.72 (0.55-0.94)	0.017	0.05
LIMK1	LIM domain kinase 1	Regulation of actin filament dynamics	Lung cancer (Chen, 2013) Alzheimer's (Heredia, 2006)	1.28 (1.06, 1.55)	0.01	0.75 (0.61-0.91)	0.0057	0.03

SIRT7	sirtuin (silent mating type information regulation 2 homolog) 7 (S. cerevisiae)	Transcription repressor	CVD (Vakhrusheva, 2007)	0.89 (0.82, 0.96)	0.0038	0.80 (0.6-1.07)	0.14	0.22
ARHGAP4	Rho GTPase Activating protein 4	regulation of small GTP-binding proteins from the RAS superfamily	cognition (Huang, 2012)	0.88 (0.81, 0.95)	0.0035	0.62 (0.38-1.00)	0.054	0.11
ATG16L2	autophagy related 16-like 2 (S. cerevisiae)	Autophagy	CVD (Magne, 2015)	0.81 (0.73, 0.90)	0.00028	0.81 (0.59-1.11)	0.19	0.25
TPM3	Tropomyosin 3	Actin-binding protein	Lung cancer (Rostila, 2012)	0.65 (0.48, 0.88)	0.0086	1.02 (0.83-1.26)	0.85	0.85
5-HTR1B	5-hydroxytryptamine (serotonin) receptor 1B	Neurotransmitter/ vasoconstriction	CVD (Iwabayashi 2012)	1.31 (1.08, 1.59)	0.0097	1.28 (0.49-3.34)	0.62	0.71
PYGO2	Pygophus homolog 2	Related to Wnt signaling	Lung cancer (Liu, 2013)	0.93 (0.85, 1.01)	0.097	0.75 (0.61-0.92)	0.0078	0.03

Models adjusted for age, BMI, SES, smoking (validation cohort), leukocyte and neutrophil count, daytime of blood sampling and season. P-values corrected for multiple testing are represented as q-values.

Legends to the figures

Figure 1. Schematic representation of the application of the modified version of the meet-in-the middle approach to identify biomarkers of disease. CVD= cardiovascular disease, CeVD= cerebrovascular disease, COPD= chronic obstructive pulmonary disease.

Figure 2: Venn diagram showing the overlap of all genes significantly associated with long-term PM10 and PM2.5 exposure in men and women in the discovery cohort .

Figure 3. Receiver operating characteristics (ROC) curve for leukocyte gene expression of gene sets distinguishing between high and low long-term PM_{10} or $PM_{2.5}$ exposure respectively, based on the 8 genes selected for validation for each gender. (A) performance of geneset consisting of *DNAJB5*, *RAC3*, *SLA2, HDLBP, PRG2, PER1, PIK3R1,* and *EAPP* to dinstinguish between high and low PM_{10} exposure in men (*above 75th percentile corresponding to: 24.5 µg/m³) and low (<24.5 µg/m³)* and (B) performance of geneset consisting of *ARHGAP4*, *AKAP6, PYGO2, HTR1B, ATG16L2, SIRT7, TPM3* and *LIMK1* in women to distinguish between high (above 75th percentile corresponding to: 25.7 µg/m³) and low (<25.7 µg/m³) long-term residential PM_{10} exposure. (C) performance of same male-specific geneset in men and (D) and female-specific geneset in women to distinguish between high (above 75th percentile corresponding to: 16.0 µg/m³) and low (<16.0 µg/m³) long-term residential $PM_{2.5}$ exposure.

Figure 1.

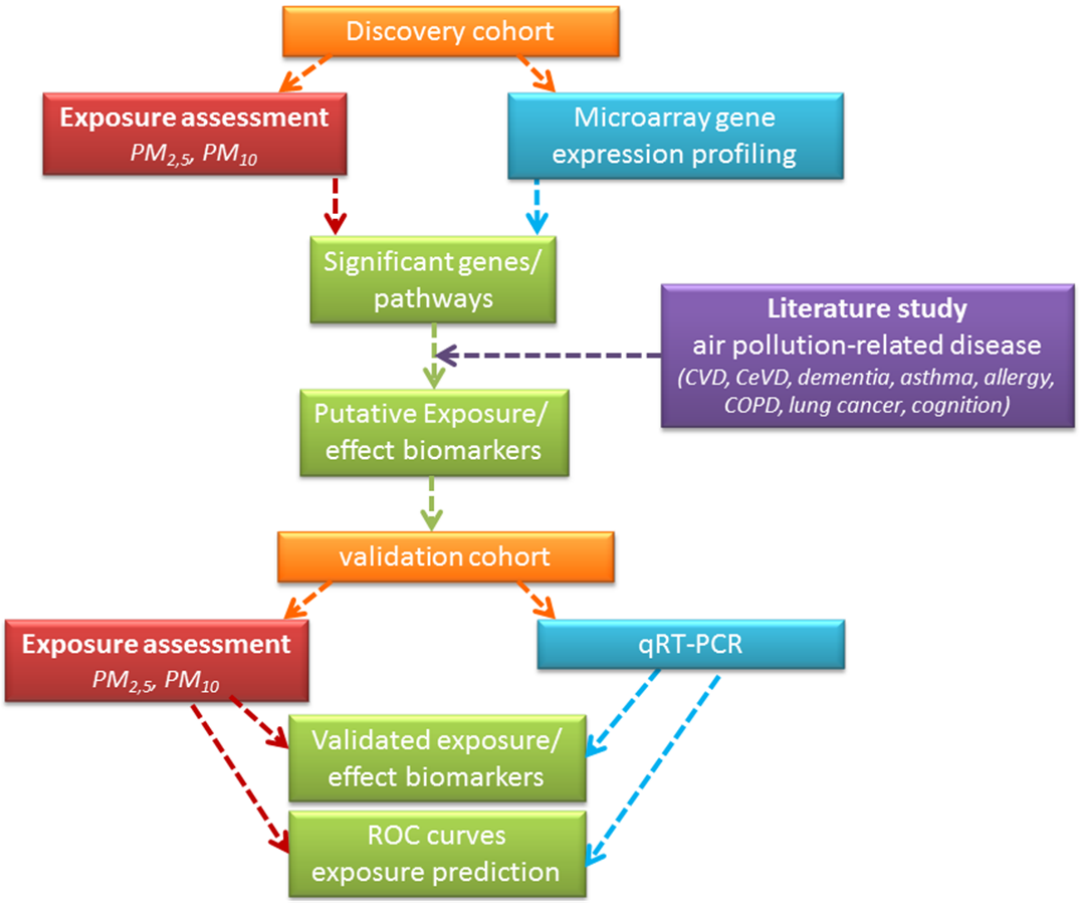

Discovery cohort
Exposure assessment
$PM_{2,5}$, PM_{10}
Microarray gene expression profiling
Significant genes/ pathways
Literature study
air pollution-related disease
(CVD, CeVD, dementia, asthma, allergy, COPD, lung cancer, cognition)
Putative Exposure/ effect biomarkers
validation cohort
Exposure assessment
$PM_{2,5}$, PM_{10}
qRT-PCR
Validated exposure/ effect biomarkers
ROC curves exposure prediction

Figure 2.

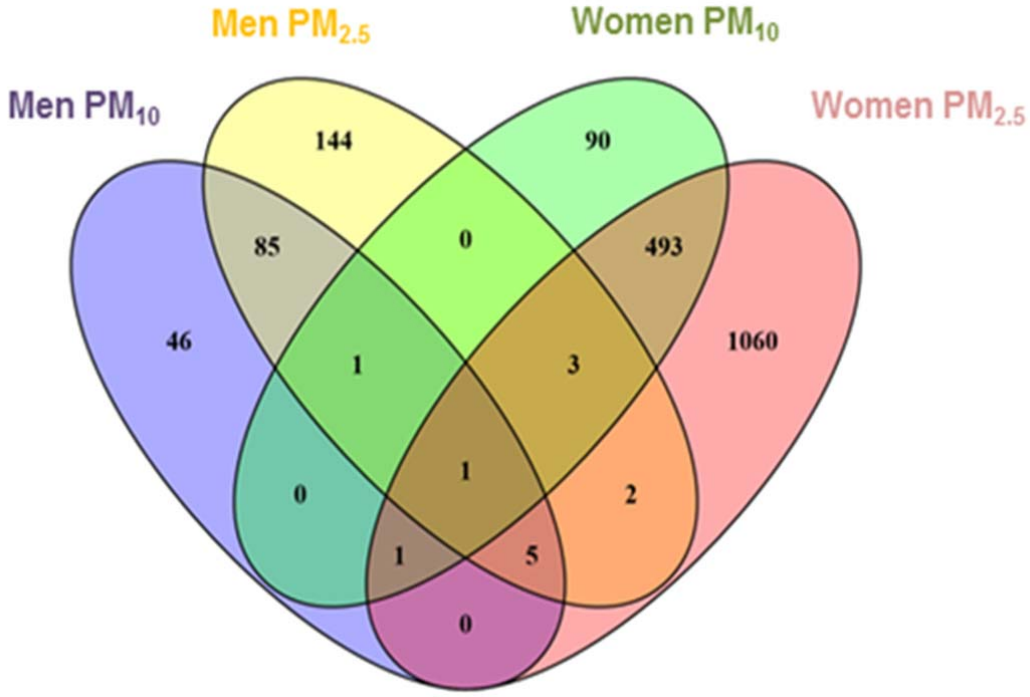
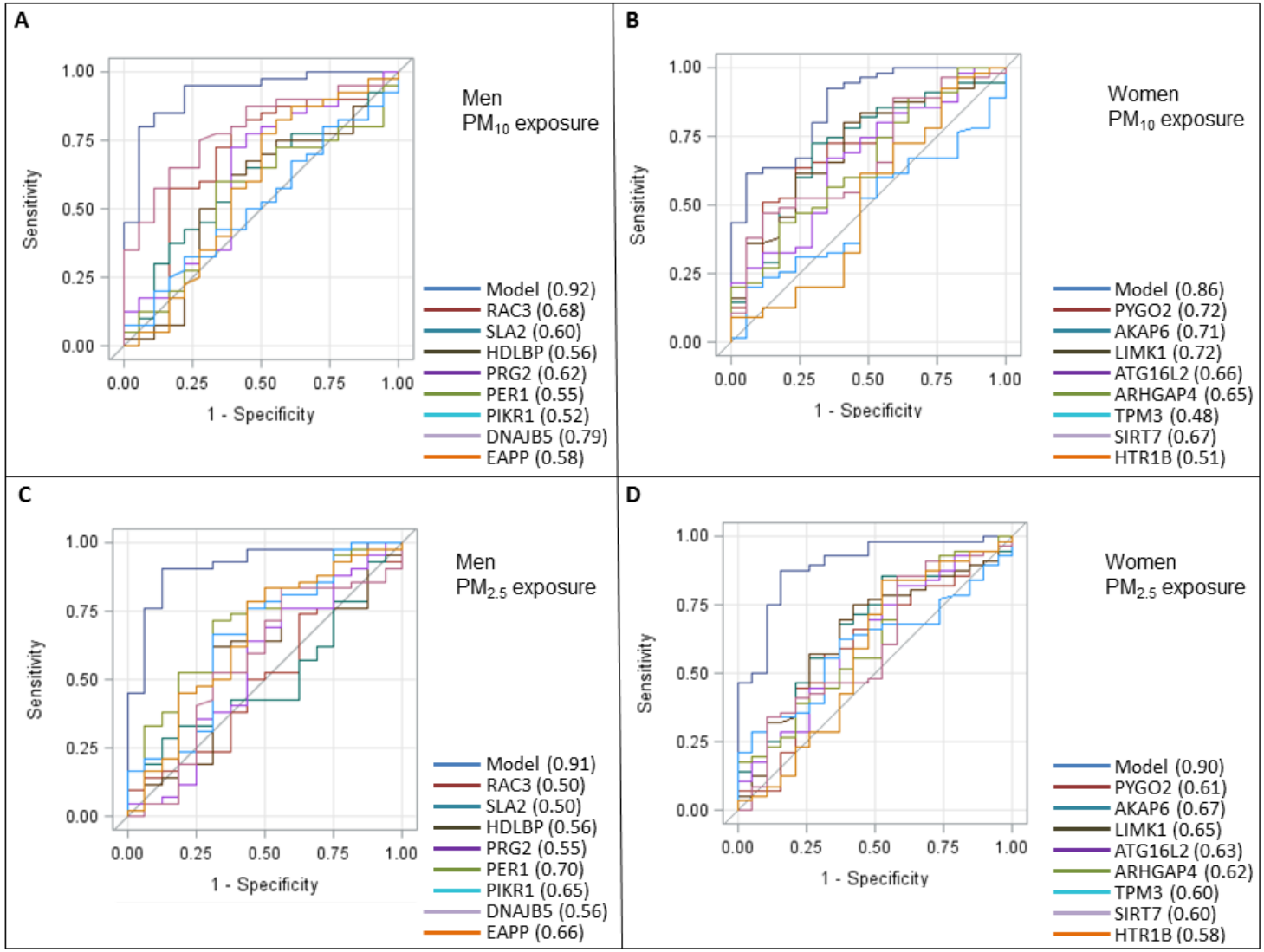

Men $PM_{2.5}$
Women PM_{10}
Men PM_{10}
Women $PM_{2.5}$
144
90
85
0
493
46
1
3
1060
1
0
2
1
5
0

Figure 3.


A
Men
PM10 exposure
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00
Model (0.92)
RAC3 (0.68)
SLA2 (0.60)
HDLBP (0.56)
PRG2 (0.62)
PER1 (0.55)
PIKR1 (0.52)
DNAJB5 (0.79)
EAPP (0.58)
B
Women
PM10 exposure
Model (0.86)
PYGO2 (0.72)
AKAP6 (0.71)
LIMK1 (0.72)
ATG16L2 (0.66)
ARHGAP4 (0.65)
TPM3 (0.48)
SIRT7 (0.67)
HTR1B (0.51)
C
Men
PM2.5 exposure
Model (0.91)
RAC3 (0.50)
SLA2 (0.50)
HDLBP (0.56)
PRG2 (0.55)
PER1 (0.70)
PIKR1 (0.65)
DNAJB5 (0.56)
EAPP (0.66)
D
Women
PM2.5 exposure
Model (0.90)
PYGO2 (0.61)
AKAP6 (0.67)
LIMK1 (0.65)
ATG16L2 (0.63)
ARHGAP4 (0.62)
TPM3 (0.60)
SIRT7 (0.60)
HTR1B (0.58)